Measurement of the $t\bar{t}$ Production Cross Section at DØ using b-tagging

1. Introduction
2. Top Quark Production and Decay
3. The lepton + jets Channel
4. Method Overview
5. Jet Tagging Efficiencies
6. Background Estimate
7. Results and Summary

Gustavo Otero y Garzón, University of Illinois at Chicago for the DØ experiment

Joint Meeting of Pacific Region Particle Physics Communities
October 29 – November 03, 2006; Honolulu, Hawaii
The top quark was discovered 11 years ago and so far it has been only observed at Fermilab

- RunI measured $\sigma_{tt\bar{t}}$ to a 25% precision with $L \sim 100\text{pb}^{-1}$ and $\sqrt{s} = 1.8 \text{ TeV}$
- At present, 30% higher production rate at $\sqrt{s} = 1.96 \text{ TeV}$ and higher luminosity

The theoretical prediction of $\sigma_{tt\bar{t}}$ has a 9-12% accuracy

Top pair events are an important background source for other physics processes
- Single Top
- Higgs search

Measuring the top pair production cross section is the first step toward any Top property analysis
Top Quark Pair Production and Decay

- Top quarks are mainly produced in pairs (strong interactions) at Tevatron energies

\[ \sigma_{\text{inel}} / \sigma_{ttbar} \sim 10^{10} \]

- High luminosity
- High efficiency

- No hadronic bound state due to short lifetime
- Electroweak decay

Final state determined by the decay of the W boson

- Dilepton channel (low bkg)
- Lepton + jets channel (moderate bkg)
- All hadronic channel (huge bkg)

Lepton = e, \( \mu \) from W or from \( \tau \) from W
The Lepton + Jets Channel

**Signal**
- 1 isolated high p_T lepton (µ, e)
- 1 ν (reconstructed as missing transverse energy (MET))
- ≥ 3 high p_T central jets

**Background**
- W (→ l ν) + ≥ 3 jets
- QCD Multijet
  - fake isolated lepton
  - misreconstructed MET
Method Overview

QCD Multijet events dominate in pp collisions

- Select a leptonically decaying W boson in association with jets
- Select objects in the final state with high efficiency and purity minimizing the instrumental background (QCD and W-like events determined from data)

Extract \( \sigma_{t\bar{t}} \) from the excess over the predicted background

Events with \( \geq 1 \) b-jet (tt has 2 b-jets!)

b-tagging

loose preselection

tight preselection
b-tagging

- $b$-quarks hadronize into long lived (ct ~ 450mm) $B$ hadrons which travel a few millimeters before decaying

- $b$-jets can be identified!
  - Soft Lepton Tagging: lepton within a jet from a semileptonic $B$ decay
  - Secondary Vertex Tagging: reconstruct SV from tracks significantly displaced from the PV originating from lifetime effects of $B$ hadrons

- To decouple tagging efficiency from tracking inefficiencies and calorimeter noise problems the SVT tagging probability is split into:
  - Probability for a jet to be taggable
  - Probability for a jet to be tagged
Taggability

- A jet is taggable if it has a matched track-jet
- Tracks in a track-jet must have hits in the SMT
  - Strong dependence on the SMT geometry

- Parameterized vs. $\eta_{\text{jet}}$ and $p_T^{\text{jet}}$ in 6 bins of $\text{sign}(P_{\text{VZ}} \times \eta_{\text{jet}}) \times |P_{\text{VZ}}|$
- Measured in $l+\text{jets}$ data

- Taggability is jet flavor dependent
- Corrected with ratios of $b(c) / \text{light taggabilities measured in MC}$
**b-Tagging Efficiency**

- Semileptonic $b$-tagging rate measured purely from 2 $\mu$-in-jet **data** samples (subset enriched in heavy flavor)

- Parameterized in terms of $p_T^{jet}$ and $\eta^{jet}$

- Calibrated by data-to-MC scale factor given by the ratio of semileptonic $b$-tagging efficiencies measured in data and $b\bar{b}$ MC

- Inclusive $b$ ($c$) tagging rates measured in $t\bar{t}$ MC and corrected with $SF_{data-MC}$

\[
SF_{data-to-MC}(p_T, \eta) = \frac{\epsilon_{b \rightarrow \mu}^{data}(p_T, \eta)}{\epsilon_{b \rightarrow \mu}^{MC}(p_T, \eta)}
\]
Mistag Rate

- A light \((u,d,s,g)\) jet identified as heavy flavor is a mistag (fake tag)
- Originated from misreconstruction and resolution effects
- Determined from the negative tagging rate
  - Measured in QCD data, dominated by light jets \((\varepsilon_{\text{data}})\)

- MC based corrections to \(\varepsilon_{\text{data}}\) for:
  - Heavy flavor contamination in QCD data (~0.5)
  - Long lived particles \(\left(K^0_S, \Lambda^0\right)\) not present in \(\varepsilon_{\text{data}}\) (~1.6)
Event Tagging Probability

Jet Taggability\_\text{flavor} \alpha 
Jet Tagging Probability\_\text{flavor} \alpha → Per Jet Probability (Weight all jets)

\alpha = b, c, light

\[ P = 1 \text{ tag} \quad \ttbar 4 \geq \text{jets} = 0.45 \]

\[ P \geq 2 \text{ tag} \quad \ttbar 4 \geq \text{jets} = 0.14 \]

- W+jets tagging probabilities:
  - Add probabilities for different flavor configurations
  - Weight each configuration with the corresponding flavor fraction
  - Fractions determined as cross section ratios of hadron-matched-jets ALPGEN MC samples

\[ \text{Trigger weight} \]
Background Estimate

- QCD-multijet production in the tagged sample:
  - Apply the Matrix Method to the tagged data sample

- Additional low rate electroweak background processes
  \[
  N^{\text{presel}}_i = \sigma^{\text{theory}}_i \times \varepsilon^{\text{presel}}_i \times \text{BR} \times L
  \]
  \[
  i = \text{single top, diboson and } Z\rightarrow\tau\tau
  \]

  \[
  N^{\text{tag\_bkg}}_i = N^{\text{presel}}_i \times P^{\text{tag}}_i
  \]

- W+jets (dominant background)
  - Overall normalization from data
  \[
  N^{\text{Presel}}_{W+jets} = N^{\text{sig}} - N^{\text{Presel}}_{tt} - \sum N^{\text{presel}}_i
  \]

  - Apply different tagging probabilities based on the flavor configurations
Cross Section Extraction

- The \( \text{ttbar} \) cross-section is calculated combining 8 channels
  - \( e + \text{jets} \) and \( \mu + \text{jets} \)
  - single tags and double tags
  - \( \geq 3 \text{jets} \) and \( \geq 4 \text{jets} \)

**Systematic Uncertainties**

<table>
<thead>
<tr>
<th>Source</th>
<th>( \sigma^+ )</th>
<th>( \sigma^- )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muon trigger</td>
<td>0.05</td>
<td>0.07</td>
</tr>
<tr>
<td>EM trigger</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>Jet trigger</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>Muon preselection</td>
<td>0.16</td>
<td>0.14</td>
</tr>
<tr>
<td>Electron preselection</td>
<td>0.17</td>
<td>0.15</td>
</tr>
<tr>
<td>Jet preselection</td>
<td>0.13</td>
<td>0.11</td>
</tr>
<tr>
<td>Preselection efficiency (MC statistics)</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td>( \varepsilon_{\text{QCD}} ) and ( \varepsilon_{\text{sig}} ) in ( \mu + \text{jets} ) channel</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>( \varepsilon_{\text{QCD}} ) and ( \varepsilon_{\text{sig}} ) in ( e + \text{jets} ) channel</td>
<td>0.06</td>
<td>0.00</td>
</tr>
<tr>
<td>Matrix Method (data statistics)</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>Taggability in data</td>
<td>0.03</td>
<td>0.00</td>
</tr>
<tr>
<td>Flavor dependence of taggability</td>
<td>0.00</td>
<td>0.03</td>
</tr>
<tr>
<td>Semileptonic ( b ) tagging efficiency in data</td>
<td>0.33</td>
<td>0.24</td>
</tr>
<tr>
<td>Semileptonic ( b ) tagging efficiency in MC</td>
<td>0.17</td>
<td>0.04</td>
</tr>
<tr>
<td>Inclusive ( b ) tagging efficiency in MC</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Inclusive ( c ) tagging efficiency in MC</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>Negative tagging efficiency in data</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>( S_{\text{BR}} ) and ( S_{\text{FMR}} )</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>( W ) fractions</td>
<td>0.29</td>
<td>0.27</td>
</tr>
<tr>
<td>( W ) fractions (MC statistics)</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>Total systematics (quad sum of the above)</td>
<td>0.57</td>
<td>0.47</td>
</tr>
<tr>
<td>Total uncertainty (nuisance parameter hold)</td>
<td>0.94</td>
<td>0.86</td>
</tr>
</tbody>
</table>
The Soft Lepton Tag Analysis

- SLT tagging depends on the physics process
  - Tagging rates measured in MC for each process (in agreement with tagging rates measured in data!)
- Same philosophy as the SVT analysis
  - Same preselected data set
  - Same backgrounds + $Z \rightarrow \mu^+\mu^-$ (determined from MC, normalized to data)
  - Similar cross section extraction procedure (single and double together)

<table>
<thead>
<tr>
<th>Source</th>
<th>$\sigma_+$ (pb)</th>
<th>$\sigma_-$ (pb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muon preselections</td>
<td>+0.18</td>
<td>-0.13</td>
</tr>
<tr>
<td>Electron preselections</td>
<td>+0.19</td>
<td>-0.13</td>
</tr>
<tr>
<td>EM triggers</td>
<td>+0.00</td>
<td>-0.03</td>
</tr>
<tr>
<td>Muon triggers</td>
<td>+0.12</td>
<td>-0.09</td>
</tr>
<tr>
<td>Jet triggers</td>
<td>+0.00</td>
<td>-0.04</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>+0.19</td>
<td>-0.12</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>+0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td>Jet ID</td>
<td>+0.14</td>
<td>-0.12</td>
</tr>
<tr>
<td>$Z + jets$ normalization</td>
<td>+0.06</td>
<td>-0.07</td>
</tr>
<tr>
<td>Heavy flavor tagging</td>
<td>+0.24</td>
<td>-0.17</td>
</tr>
<tr>
<td>Fake tagging rate</td>
<td>+0.84</td>
<td>-0.78</td>
</tr>
<tr>
<td>Matrix method</td>
<td>+0.33</td>
<td>-0.35</td>
</tr>
<tr>
<td>Monte Carlo statistics</td>
<td>+0.25</td>
<td>-0.27</td>
</tr>
<tr>
<td>$W$ fractions</td>
<td>+0.13</td>
<td>-0.19</td>
</tr>
<tr>
<td>Total systematics (quad sum of the above)</td>
<td>+1.04</td>
<td>-0.98</td>
</tr>
</tbody>
</table>
Perform a maximum likelihood fit to the observed number of events incorporating all systematic uncertainties in the fit using a nuisance parameter likelihood method.

\[ \sigma_{\text{ttbar}} = 6.6 \pm 0.9(\text{stat+syst}) \pm 0.4(\text{lum}) \text{ pb} \]  

\[ \sigma_{\text{ttbar}} = 7.3^{+2.0}_{-1.8}(\text{stat+syst}) \pm 0.4(\text{lum}) \text{ pb} \]

Based on 425 pb\(^{-1}\) of DØ data.

\[ \sigma_{\text{ttbar}}^{\text{NLO}} = 6.8 \pm 0.6 \text{ pb} \]

Summary

- Entering the top quark precision measurement era at the Tevatron
- Results in agreement with the Standard Model prediction at NLO
- First preliminary DØ RunII result with SLT
- SVT analysis currently DØ’s most precise top cross section measurement (down to 15% uncertainty)
  - Better understanding of systematic uncertainties (12% stat ⊕ 8% syst)
  - At the DØ gate on its way to PRD (FERMILAB-PUB-06-386-E)

... expect fb⁻¹ results soon!
Back Up Slides
**The Tevatron and DØ**

- **Tevatron**
  - Proton anti-proton collider
  - $\sqrt{s} = 1.96$ TeV
  - $\sim 1.8$ fb$^{-1}$ delivered
  - Expected x40-80 Run I data set!

- **DØ**
  - A truly international collaboration
  - Multipurpose detector
    - central tracking embedded in a solenoidal field
    - preshowers
    - EM and hadronic calorimeters
    - muon system
Why study the Top Quark?

- Predicted in the ’70s by the SM and discovered in 1995
  - Least well studied component of the SM (only produced at the Tevatron so far)

- Only known fermion with a mass at the natural Electroweak scale

- Lifetime (5x10^{-25} s) shorter than the hadronization time (no top hadronic bound states)

- The top quark is relevant for many Electroweak analyses

- Strongest coupling to the Higgs
  (Yukawa coupling \( \lambda_t \propto m_t \sim 1 \))
The Top Quark and the SM Higgs

Corrections to $W$ and $Z$ boson masses from top quark and Higgs boson loops constrain the Higgs boson mass

\[ M_W^2 = M_W^{(0)} \times (1 - \Delta) \]

\[ \Delta_t^{-1} \sim M_t^2 \]

\[ \Delta_H \sim \ln (M_H^2) \]

Goal for Tevatron in Run II:

\[ \Delta M_t = 3 \text{ GeV} \]

\[ \Delta M_W = 20 \text{ MeV} \]
Top Quark Physics

- Production Cross Section
- Production Kinematics
- Top Spin Polarization
- Resonance Production
- Top Mass
- W Helicity
- Top Charge
- Branching Ratios
- $|V_{tb}|$
- Spin Correlation
- Non-SM decays

Gustavo Otero y Garzón, UIC

October 30th, 2006
Top Quark Pair Production

Top quarks are mainly produced in pairs (through strong interactions) at Tevatron energies (electroweak production to be observed soon!)

\[ x = \frac{2 \times m_t}{\sqrt{s}} \]

<table>
<thead>
<tr>
<th>Energy</th>
<th>( \sigma_{tt}^{NLO} ) (pb)</th>
<th>( q\bar{q} \rightarrow t\bar{t} )</th>
<th>( gg \rightarrow t\bar{t} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Run I (1.8 TeV)</td>
<td>4.87±10%</td>
<td>90%</td>
<td>10%</td>
</tr>
<tr>
<td>Run II (2.0 TeV)</td>
<td>6.70±10%</td>
<td>85%</td>
<td>15%</td>
</tr>
<tr>
<td>LHC (14 TeV)</td>
<td>803±15%</td>
<td>10%</td>
<td>90%</td>
</tr>
</tbody>
</table>

\( \sigma_{inel} / \sigma_{ttbar} \sim 10^{10} \)

- High luminosity
- High efficiency
Top Quark Decay

- $\Gamma_t^{SM} \approx 1.5 \text{ GeV } (m_t=175 \text{ GeV}) \Rightarrow \tau \sim 10^{-25} \text{ s } \Rightarrow \text{no hadronic bound states}$

- Top quark decays via the Weak interaction exclusively as $t \rightarrow W b$
  - $|V_{tb}| > 0.999$
  - $R = 1.03 \pm 0.19$ (hep-ex/0503002)
  - Negligible rates for FCNC ($t \rightarrow q \gamma, Z, g$)

- Final state determined by the decay of the W boson:
  - dilepton channel (low bkg)
  - lepton + jets channel (moderate bkg)
  - all hadronic channel (huge bkg)

Lepton$= e, \mu$ from $W$ or $\tau$ from $W$
Preselected Sample

- Require:
  - All events pass the signal trigger
  - A tight isolated lepton
  - Large MET (neutrino)
  - At least one jet
  - MET separated from the lepton in the transverse plane
  - Second lepton veto (orthogonal to dilepton analyses)

- Composition of the preselected sample determined by defining two samples (Matrix Method)
  - Tight sample : the preselected sample \( N_t \)
  - Loose sample : events passing the preselection but with a loose lepton requirement

\[
N_t = N_{\text{sig}} + N_{\text{QCD}}
\]

\[
\varepsilon_{\text{sig}} \downarrow \quad \varepsilon_{\text{QCD}} \downarrow
\]

\[
N_t = \varepsilon_{\text{sig}} N_{\text{sig}} + \varepsilon_{\text{QCD}} N_{\text{QCD}}
\]

- \( \varepsilon_{\text{sig}} \) → efficiency for a loose lepton from a W decay to pass the tight criteria ~85%
- \( \varepsilon_{\text{QCD}} \) → rate for a loose lepton in QCD to appear to be tight ~15%

<table>
<thead>
<tr>
<th></th>
<th>1 jet</th>
<th>2 jets</th>
<th>3 jets</th>
<th>≥ 4 jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N_t )</td>
<td>6153</td>
<td>2217</td>
<td>466</td>
<td>119</td>
</tr>
<tr>
<td>( N_{\text{sig}}^t )</td>
<td>5806±82</td>
<td>1976±58</td>
<td>395±24</td>
<td>99.8±11.7</td>
</tr>
<tr>
<td>( N_{\text{QCD}}^t )</td>
<td>347±14</td>
<td>241±11</td>
<td>71±5</td>
<td>19.2±2.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1 jet</th>
<th>2 jets</th>
<th>3 jets</th>
<th>≥ 4 jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N_t )</td>
<td>6827</td>
<td>2267</td>
<td>439</td>
<td>100</td>
</tr>
<tr>
<td>( N_{\text{sig}}^t )</td>
<td>6607±85</td>
<td>2155±50</td>
<td>406±22</td>
<td>91.4±10.7</td>
</tr>
<tr>
<td>( N_{\text{QCD}}^t )</td>
<td>220±12</td>
<td>112±10</td>
<td>33±5</td>
<td>8.6±2.1</td>
</tr>
</tbody>
</table>

- Top events populate 3rd and 4th jet multiplicity bins
- Events with 1 and 2 jets used as control of background estimate
Event Simulation

- ttbar and W+jets are generated using ALPGEN 1.3 followed by PYTHIA 6.2 to simulate the underlying event and the hadronization

- Other small backgrounds
  - Single top (CompHEP/PYTHIA)
  - Diboson (ALPGEN/PYTHIA) + NLO K-factor
  - Z/gamma* (PYTHIA) @ NNLO

- TAUOLA simulates \( \tau \) decays

- W+jets samples are generated separately for processes with 1, 2, 3, and 4 or more partons in the final state using ALPGEN

<table>
<thead>
<tr>
<th>process</th>
<th>( \sigma ) (pb)</th>
<th>NLO correction</th>
<th>Branching ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>( tb \to \ell \nu bb )</td>
<td>0.88</td>
<td>-</td>
<td>0.1259</td>
</tr>
<tr>
<td>( tbq \to \ell \nu bbj )</td>
<td>1.98</td>
<td>-</td>
<td>0.1259</td>
</tr>
<tr>
<td>( WW \to \ell \nu jj )</td>
<td>2.04</td>
<td>1.31</td>
<td>0.3928</td>
</tr>
<tr>
<td>( WZ \to \ell \nu jj )</td>
<td>0.61</td>
<td>1.35</td>
<td>0.3928</td>
</tr>
<tr>
<td>( WZ \to jj \ell \ell )</td>
<td>0.18</td>
<td>1.35</td>
<td>0.4417</td>
</tr>
<tr>
<td>( ZZ \to jj \ell \ell )</td>
<td>0.16</td>
<td>1.28</td>
<td>0.4417</td>
</tr>
<tr>
<td>( Z/\gamma^* \to \tau \tau )</td>
<td>253</td>
<td>-</td>
<td>0.3250</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>process ( \sigma ) (pb)</th>
<th>process ( \sigma ) (pb)</th>
<th>process ( \sigma ) (pb)</th>
<th>process ( \sigma ) (pb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( W_j ) 1600</td>
<td>( W_j ) 517</td>
<td>( W_j ) 163</td>
<td>( W_{jjj} ) 49.5</td>
</tr>
<tr>
<td>( W_c ) 51.8</td>
<td>( W_{cj} ) 28.6</td>
<td>( W_{cj} ) 19.4</td>
<td>( W_{ccJ} ) 3.15</td>
</tr>
<tr>
<td>( W_{bb} ) 9.85</td>
<td>( W_{bbJ} ) 5.24</td>
<td>( W_{bbJ} ) 2.86</td>
<td>( W_{c\ell} ) 5.83</td>
</tr>
<tr>
<td>( W_{cc} ) 24.3</td>
<td>( W_{ccJ} ) 12.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE XIII:** W+jets boson processes in ALPGEN and their cross sections for the leptonic W boson decay, \( \sigma \equiv \sigma_{pp \to W+j, \ell} Br(W \to ll) \), where \( j = u, d, s, g \) and \( J = u, d, s, g, c \).
W+jets event generation

- LO parton level calculations from ALPGEN need to be combined with the partonic evolution from PYTHIA to avoid double counting of configurations leading to the same final state.
- Ad-hoc MLM matching is used: matrix element partons are matched to reconstructed jets within a 0.5 cone and classified according to the number of HF jets in the final state.
  - Keep events only if the number of reconstructed jets equals the number of matrix element partons in the 1,2,3 jets bin.
  - Keep all events with >=4 reconstructed jets in the n>=4 jets bin, independently of the additional number of non-matched light jets.
- NLO K-factor applied to the Wbbar, Wccbar, W(bbbar) and W(ccbar)

<table>
<thead>
<tr>
<th>Contribution</th>
<th>W+1 jet</th>
<th>W+2 jets</th>
<th>W+3 jets</th>
<th>W+4 jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wbbar</td>
<td>(1.23 ± 0.08)%</td>
<td>(2.05 ± 0.21)%</td>
<td>(2.84 ± 0.16)%</td>
<td></td>
</tr>
<tr>
<td>Wccbar</td>
<td>(1.69 ± 0.12)%</td>
<td>(2.94 ± 0.37)%</td>
<td>(4.44 ± 0.29)%</td>
<td></td>
</tr>
<tr>
<td>W(bb)</td>
<td>(0.86 ± 0.03)%</td>
<td>(1.46 ± 0.09)%</td>
<td>(2.03 ± 0.15)%</td>
<td>(2.99 ± 0.24)%</td>
</tr>
<tr>
<td>W(cc)</td>
<td>(1.23 ± 0.05)%</td>
<td>(2.26 ± 0.15)%</td>
<td>(3.08 ± 0.24)%</td>
<td>(5.06 ± 0.54)%</td>
</tr>
<tr>
<td>Wc</td>
<td>(4.41 ± 0.18)%</td>
<td>(6.25 ± 0.43)%</td>
<td>(4.93 ± 0.48)%</td>
<td>(4.30 ± 0.23)%</td>
</tr>
<tr>
<td>W+light</td>
<td>(93.5 ± 0.2)%</td>
<td>(87.1 ± 0.7)%</td>
<td>(85.0 ± 1.1)%</td>
<td>(80.4 ± 0.7)%</td>
</tr>
</tbody>
</table>

Systematic error on the fractions comes from difference between matching schemes, choice of matching cone size, PDF and renormalization and factorization scales.
Jet Flavor Identification in MC

- For all MC samples, the jet flavor (b, c, or light) is determined by matching the direction of the reconstructed jet to the hadron flavor within a cone of $R=0.5$

- If more than one hadron is found within the cone, the jet is considered a:
  - b-jet, if the cone contains at least one B hadron
  - c-jet, if the cone contains at least one C, and no B hadron
  - light, if the cone contains no B or C hadron.
**W+jets Background**

- The dominant background by far
- Overall normalization before tagging obtained directly from Data

\[ N_{\text{presel}}^{(W \rightarrow l_\nu)+n_j} = N_{t}^{\text{sig}} - N_{tt\rightarrow l+jets}^{\text{presel}} - N_{tt\rightarrow ll}^{\text{presel}} - \sum_{bkg i} N_{bkg i}^{\text{presel}} \]

- Number of W+jets events in the tagged sample determined by

\[ N_{\text{tag}}^{(W \rightarrow l)+n_j} = N_{\text{presel}}^{(W \rightarrow l)+n_j} P_{\text{tag}}^{(W \rightarrow l)+n_j} \]

- The event tagging probability results from adding the tagging probabilities for the different flavor configurations weighted with their fractions

\[ P_{\text{tag}}^{(W \rightarrow l)+n_j} = \sum_{\Phi_n} F_{\Phi_n} P_{\Phi_n}^{\text{tag}} \]

\[ F_{\Phi,n} = \frac{\sigma_{\Phi,n}^{\text{eff}}}{\sum_{\Phi} \sigma_{\Phi,n}^{\text{eff}}} \]
Secondary Vertex Tagger Algorithm

The Secondary Vertex Tagger (SVT) is a lifetime tagger that explicitly reconstructs vertices which are displaced from the Primary Vertex

- Use of tracks with significant impact parameter with respect to the Primary Vertex

- Build-up method fitting pairs of selected tracks within track-jets

- Removes track pairs in the mass windows corresponding to $K^0_S$, $\Lambda^0$ and photon conversions ($\gamma \rightarrow e^+e^-$)

- A jet is identified as a b-jet (tagged) if it contains a reconstructed secondary vertex within a jet

Top events have two b-jets while events from other processes very seldom have heavy flavor!
Data Sample

- Run Quality, Luminosity Block and Event Quality Selections applied
- A lepton and a jet are required at trigger level
- Trigger efficiencies are estimated by folding into the MC the per-object individual trigger conditions measured in Data

![Graph showing Run II Integrated Luminosity from 19 April 2002 to 8 October 2006]

- **Data set used**
Analysis Method

- Determine the number of selected events for each background
- Parameterize the tagging efficiencies determined in data
- Determine event tagging probabilities for all the backgrounds
- Use the Monte Carlo simulation event kinematics and fold in the tagging efficiencies from data to estimate the number of tagged events
- Estimate the ttbar cross section from the excess in the actual number of tagged events with 3 and ≥ 4 jets over the background prediction

\[
\sigma = \frac{N_{\text{tag observed}} - N_{\text{tag background}}}{BR \cdot L \cdot \varepsilon_{\text{presel}} \cdot \bar{P}_{\text{tag}}}
\]
Semileptonic $b$-Tagging Efficiency

- $b$-tagging rate is measured purely from data applying SVT and Soft Lepton Tagger to two samples:
  - muon-in-jet
  - muon-in-jet away jet tagged (enriched in heavy flavor)

- Use system8
  - Samples with different fractions of signal and background
  - SVT and SLT have different efficiencies for signal and background
  - SVT and SLT decorrelated
  - 8 equations with 8 unknowns

- Parameterize in terms of $p_T^{jet}$ and $y^{jet}$

Systematic uncertainties arise from the variation of the correlation parameters in System8
Inclusive Tagging Efficiencies

- Inclusive $b(c)$-tagging efficiency in MC
  - Measured in $\bar{t}t$bar

  **b-inclusive ($\bar{t}t$bar MC)**

  **c-inclusive ($\bar{t}t$bar MC)**

- Calibrated by data-to-MC scale factor given by the ratio of semileptonic $b$-tagging efficiencies measured in data and $bb\bar{b}$ MC

\[
SF_{b\to\mu}(p_T, \eta) = \frac{\epsilon_{b\to\mu}^{data}(p_T, \eta)}{\epsilon_{b\to\mu}^{MC}(p_T, \eta)}
\]

Systematic uncertainties on the inclusive efficiencies coming from the difference between parameterizations obtained in $\bar{t}t$bar MC with two choices of $b$-fragmentation models. Systematic uncertainty on semileptonic $b$ efficiency in MC takes as the difference in efficiency between $bb\bar{b}$ and $\bar{t}t$bar.
Mistag Rate

The Negative Tagging Rate is corrected for:

- Heavy flavor contamination in QCD data (estimated in QCD MC)

\[
SF_{hf}(p_T, y) = \frac{\varepsilon^\text{light}(p_T, y)}{\varepsilon^\text{inclusive}(p_T, y)}
\]

- Remaining long lived particles ($K^0_s$, $\Lambda^0$) not present in the negative tagging rate (estimated in QCD MC)

\[
SF_{ll}(p_T, y) = \frac{\varepsilon^\text{+ light}(p_T, y)}{\varepsilon^\text{- light}(p_T, y)}
\]

- The mistag rate is then

\[
\varepsilon^\text{+ light}(p_T, y) = \varepsilon^\text{data}(p_T, y)SF_{hf}(p_T, y)SF_{ll}(p_T, y)
\]

Systematic uncertainties determined by varying by 20% the b and c fractions in PYTHIA QCD MC used for the SFs
Composition of the Preselected Sample

- The Matrix Method separates QCD from Physics Backgrounds

- Expected number of signal events:
  \[ N_{\text{presel}}^{\text{ttbar}} = \sigma_{\text{theory}}^{\text{ttbar}} \times \varepsilon_{\text{presel}}^{\text{ttbar}} \times \text{BR} \times L \]

- Expected number of non-\(W\) background events:
  \[ N_{\text{presel}}^{i} = \sigma_{\text{theory}}^{i} \times \varepsilon_{\text{presel}}^{i} \times \text{BR} \times L \]
  \(i = \text{single top, diboson (WW, WZ, ZZ) and } Z \rightarrow \tau\tau\)

<table>
<thead>
<tr>
<th>Event Category</th>
<th>1 jet [ \pm ]</th>
<th>2 jets [ \pm ]</th>
<th>3 jets [ \pm ]</th>
<th>(\geq 4) jets [ \pm ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>(tt \rightarrow t^{+}t^{-})</td>
<td>0.770 (\pm) 0.029</td>
<td>5.29 (\pm) 0.07</td>
<td>11.89 (\pm) 0.11</td>
<td>9.59 (\pm) 0.10</td>
</tr>
<tr>
<td>(tt \rightarrow t\bar{t})</td>
<td>4.04 (\pm) 0.07</td>
<td>11.55 (\pm) 0.11</td>
<td>4.21 (\pm) 0.07</td>
<td>0.667 (\pm) 0.029</td>
</tr>
<tr>
<td>(tb)</td>
<td>5.96 (\pm) 0.12</td>
<td>13.21 (\pm) 0.17</td>
<td>2.27 (\pm) 0.07</td>
<td>0.212 (\pm) 0.023</td>
</tr>
<tr>
<td>(tq)</td>
<td>5.38 (\pm) 0.11</td>
<td>10.82 (\pm) 0.15</td>
<td>3.76 (\pm) 0.09</td>
<td>0.775 (\pm) 0.044</td>
</tr>
<tr>
<td>(WW \rightarrow l\nu j j)</td>
<td>6.37 (\pm) 0.23</td>
<td>7.06 (\pm) 0.24</td>
<td>0.461 (\pm) 0.064</td>
<td>0.000 (\pm) 0.000</td>
</tr>
<tr>
<td>(WZ \rightarrow l\nu j j)</td>
<td>5.64 (\pm) 0.21</td>
<td>7.92 (\pm) 0.25</td>
<td>0.565 (\pm) 0.071</td>
<td>0.061 (\pm) 0.023</td>
</tr>
<tr>
<td>(WZ \rightarrow j j l l)</td>
<td>0.601 (\pm) 0.065</td>
<td>0.840 (\pm) 0.078</td>
<td>0.308 (\pm) 0.047</td>
<td>0.006 (\pm) 0.006</td>
</tr>
<tr>
<td>(Z Z \rightarrow j j l l)</td>
<td>0.850 (\pm) 0.071</td>
<td>1.09 (\pm) 0.08</td>
<td>0.296 (\pm) 0.043</td>
<td>0.037 (\pm) 0.015</td>
</tr>
<tr>
<td>(Z \rightarrow \tau^{+}\tau^{-})</td>
<td>0.025 (\pm) 0.002</td>
<td>0.012 (\pm) 0.002</td>
<td>0.003 (\pm) 0.001</td>
<td>0.001 (\pm) 0.000</td>
</tr>
</tbody>
</table>

- Expected number of \(W\) background events:
  \[ N_{\text{presel}}^{W} = N_{\text{sig}} - N_{\text{presel}}^{\text{ttbar}} - \sum N_{\text{presel}}^{i} \]

Preselection Efficiencies in the \(e+jets\) channel

Gustavo Otero y Garzón, UIC
October 30\textsuperscript{th}, 2006
Lepton+Jets Single Tag 3rd Jet Bin
Estimate the $tt\bar{t}$ cross-section from observed excess in the number of tagged events with respect to the background prediction

Optimum use of the statistical information (single and double tags)
## Observed vs. Predicted single tags

<table>
<thead>
<tr>
<th></th>
<th>1 jet</th>
<th>2 jets</th>
<th>3 jets</th>
<th>≥4 jets</th>
<th>1 jet</th>
<th>2 jets</th>
<th>3 jets</th>
<th>≥4 jets</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>$W + \text{light}$</strong></td>
<td>21.3±0.7</td>
<td>10.4±0.7</td>
<td>2.52±0.19</td>
<td>0.62±0.13</td>
<td>24.9±0.8</td>
<td>13.2±0.8</td>
<td>2.68±0.18</td>
<td>0.49±0.10</td>
</tr>
<tr>
<td><strong>$W (c\bar{c})$</strong></td>
<td>6.7±0.1</td>
<td>3.7±0.2</td>
<td>0.90±0.07</td>
<td>0.27±0.06</td>
<td>7.6±0.1</td>
<td>4.3±0.2</td>
<td>0.91±0.06</td>
<td>0.26±0.06</td>
</tr>
<tr>
<td><strong>$W (b\bar{b})$</strong></td>
<td>19.2±0.4</td>
<td>9.9±0.3</td>
<td>2.31±0.17</td>
<td>0.61±0.13</td>
<td>21.6±0.4</td>
<td>10.9±0.3</td>
<td>2.36±0.15</td>
<td>0.54±0.11</td>
</tr>
<tr>
<td><strong>$W_c$</strong></td>
<td>24.8±0.5</td>
<td>11.5±0.4</td>
<td>1.58±0.12</td>
<td>0.26±0.05</td>
<td>27.6±0.5</td>
<td>12.1±0.4</td>
<td>1.59±0.11</td>
<td>0.21±0.04</td>
</tr>
<tr>
<td><strong>$W_{c\bar{c}}$</strong></td>
<td>5.1±0.2</td>
<td>1.43±0.15</td>
<td>0.43±0.09</td>
<td>11.2±0.3</td>
<td>5.6±0.2</td>
<td>1.65±0.13</td>
<td>0.36±0.08</td>
<td></td>
</tr>
<tr>
<td><strong>$W_{bb}$</strong></td>
<td>10.3±0.3</td>
<td>3.08±0.23</td>
<td>0.74±0.15</td>
<td>3.10±0.21</td>
<td>11.2±0.3</td>
<td>5.6±0.2</td>
<td>1.65±0.13</td>
<td>0.36±0.08</td>
</tr>
<tr>
<td><strong>$W + \text{jets}$</strong></td>
<td>72.0±0.9</td>
<td>50.9±1.0</td>
<td>11.8±0.4</td>
<td>2.93±0.26</td>
<td>81.7±1.0</td>
<td>57.3±1.0</td>
<td>12.3±0.4</td>
<td>2.49±0.22</td>
</tr>
<tr>
<td><strong>QCD</strong></td>
<td>7.1±1.5</td>
<td>10.0±1.7</td>
<td>5.5±1.2</td>
<td>2.95±0.97</td>
<td>7.2±0.7</td>
<td>5.8±0.6</td>
<td>1.58±0.36</td>
<td>2.78±0.40</td>
</tr>
<tr>
<td><strong>single top</strong></td>
<td>2.84±0.06</td>
<td>6.28±0.09</td>
<td>1.62±0.05</td>
<td>0.26±0.02</td>
<td>2.29±0.04</td>
<td>5.56±0.06</td>
<td>1.48±0.04</td>
<td>0.24±0.02</td>
</tr>
<tr>
<td><strong>diboson</strong></td>
<td>1.95±0.08</td>
<td>2.4±0.1</td>
<td>0.20±0.03</td>
<td>&lt; 0.01</td>
<td>1.96±0.09</td>
<td>2.53±0.10</td>
<td>0.19±0.02</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td><em><em>$Z/\gamma^</em> \rightarrow \tau^+\tau^-$</em>*</td>
<td>0.13±0.04</td>
<td>0.34±0.06</td>
<td>0.02±0.01</td>
<td>&lt; 0.01</td>
<td>0.16±0.06</td>
<td>0.25±0.04</td>
<td>0.08±0.04</td>
<td>0.01±0.02</td>
</tr>
<tr>
<td>$N_{bkgd}$</td>
<td>84.0±1.7</td>
<td>69.9±1.9</td>
<td>19.1±1.3</td>
<td>6.2±1.0</td>
<td>93.3±1.2</td>
<td>71.5±1.2</td>
<td>15.6±0.6</td>
<td>5.5±0.5</td>
</tr>
<tr>
<td><strong>syst.</strong></td>
<td>+10.9-12.0</td>
<td>+8.7-9.2</td>
<td>+2.0-2.1</td>
<td>+0.5-0.5</td>
<td>+12.2-13.4</td>
<td>+9.4-10.0</td>
<td>+2.0-2.1</td>
<td>+0.5-0.5</td>
</tr>
<tr>
<td><strong>$tt \rightarrow l+\text{jets}$</strong></td>
<td>0.92±0.04</td>
<td>10.2±0.1</td>
<td>24.75±0.21</td>
<td>17.70±0.19</td>
<td>0.60±0.03</td>
<td>7.7±0.1</td>
<td>21.66±0.21</td>
<td>17.64±0.19</td>
</tr>
<tr>
<td><strong>$t\bar{t} \rightarrow ll$</strong></td>
<td>2.10±0.04</td>
<td>6.5±0.1</td>
<td>2.16±0.04</td>
<td>0.30±0.01</td>
<td>1.47±0.03</td>
<td>5.5±0.1</td>
<td>2.00±0.04</td>
<td>0.27±0.01</td>
</tr>
<tr>
<td>$N_{\text{pred}}$</td>
<td>87.0±1.7</td>
<td>86.6±1.9</td>
<td>46.0±1.3</td>
<td>24.1±1.0</td>
<td>95.4±1.2</td>
<td>84.7±1.2</td>
<td>39.3±0.6</td>
<td>23.4±0.5</td>
</tr>
<tr>
<td><strong>syst.</strong></td>
<td>+11.0-12.1</td>
<td>+8.8-9.3</td>
<td>+2.2-2.3</td>
<td>+1.4-1.4</td>
<td>+12.3-13.5</td>
<td>+9.5-10.1</td>
<td>+2.3-2.4</td>
<td>+1.3-1.4</td>
</tr>
<tr>
<td>$N_{\text{obs}}$</td>
<td>95</td>
<td>82</td>
<td>47</td>
<td>33</td>
<td>105</td>
<td>68</td>
<td>41</td>
<td>26</td>
</tr>
</tbody>
</table>
## Observed vs. Predicted double tags

<table>
<thead>
<tr>
<th></th>
<th>$e^+\text{jets}$</th>
<th>$\mu^+\text{jets}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2 jets</td>
<td>3 jets</td>
</tr>
<tr>
<td>$W+\text{light}$</td>
<td>0.017±0.003 &lt; 0.01 &lt; 0.01</td>
<td>0.027±0.003 &lt; 0.01 &lt; 0.01</td>
</tr>
<tr>
<td>$W(\bar{c}c)$</td>
<td>0.014±0.002 &lt; 0.01 &lt; 0.01</td>
<td>0.019±0.003 &lt; 0.01 &lt; 0.01</td>
</tr>
<tr>
<td>$W(bb)$</td>
<td>0.13±0.03 0.06±0.01 &lt; 0.01</td>
<td>0.29±0.05 0.05±0.01 0.02±0.01</td>
</tr>
<tr>
<td>$Wc$</td>
<td>0.027±0.002 &lt; 0.01 &lt; 0.01</td>
<td>0.039±0.003 &lt; 0.01 &lt; 0.01</td>
</tr>
<tr>
<td>$W\bar{c}\bar{c}$</td>
<td>0.24±0.01 0.07±0.01 0.02±0.01</td>
<td>0.28±0.01 0.09±0.01 0.02±0.01</td>
</tr>
<tr>
<td>$Wb\bar{b}$</td>
<td>2.80±0.13 0.86±0.08 0.22±0.05</td>
<td>3.30±0.14 0.87±0.07 0.17±0.04</td>
</tr>
<tr>
<td>$W+\text{jets}$</td>
<td>3.23±0.13 1.00±0.08 0.26±0.05</td>
<td>3.96±0.15 1.02±0.08 0.22±0.04</td>
</tr>
<tr>
<td>QCD</td>
<td>&lt; 0.01 0.27±0.02 &lt; 0.01</td>
<td>0.26±0.29 &lt; 0.01 &lt; 0.01</td>
</tr>
<tr>
<td>Single top</td>
<td>1.07±0.02 0.39±0.02 0.07±0.01</td>
<td>0.93±0.01 0.37±0.01 0.07±0.01</td>
</tr>
<tr>
<td>Diboson</td>
<td>0.34±0.02 0.04±0.01 &lt; 0.01</td>
<td>0.26±0.02 0.03±0.01 &lt; 0.01</td>
</tr>
<tr>
<td>$Z\rightarrow\tau^+\tau^-$</td>
<td>&lt; 0.01 &lt; 0.01 &lt; 0.01</td>
<td>&lt; 0.01 0.02±0.02 &lt; 0.01</td>
</tr>
<tr>
<td>$N_{\text{bkg}}$</td>
<td>4.64±0.28 1.70±0.40 0.34±0.29</td>
<td>5.42±0.33 1.44±0.34 0.29±0.38</td>
</tr>
<tr>
<td>Syst.</td>
<td>+0.83−0.81 +0.26−0.25 +0.06−0.06</td>
<td>+0.99−0.97 +0.27−0.25 +0.05−0.06</td>
</tr>
<tr>
<td>$tt\rightarrow l+p\text{jets}$</td>
<td>1.72±0.19 7.3±0.3 6.9±0.2</td>
<td>1.02±0.15 6.2±0.3 6.3±0.3</td>
</tr>
<tr>
<td>$tt\rightarrow ll$</td>
<td>1.81±0.02 0.65±0.01 0.09±0.01</td>
<td>1.50±0.02 0.61±0.01 0.08±0.01</td>
</tr>
<tr>
<td>$N_{\text{pred}}$</td>
<td>8.2±0.3 9.7±0.4 7.3±0.3</td>
<td>7.9±0.4 8.3±0.3 6.7±0.4</td>
</tr>
<tr>
<td>Syst.</td>
<td>+0.8−1.9 +0.6−1.3 +0.4−1.8</td>
<td>+1.3−1.0 +1.3−0.7 +1.7−0.4</td>
</tr>
<tr>
<td>$N_{\text{obs}}$</td>
<td>12 2 11</td>
<td>6 3 8</td>
</tr>
</tbody>
</table>